

GGCAGAACGGGCATAGTGGAACGGACCGTGTGGTATGCTCCGCGATGTCTGAGGCGACTTGGTG
GTACCGAGGAGGGACTTCAAAACATGACCTGCATTACAGAAGGGAAGCGGAGGTTAACACCACA
CTCGAGGAGTTGTTACTCTATTTTATTTTCTTAATAAATCTATGCATATTGACTTTTGGGATGG
TCAATCCACATATGTACTATTTAAACAAAGTTATGTCGTCTCTGTTTGTGGACACTTCTCTACC
TGATGATGAAAGAAGCAGCTTTAGGTCCATTCGGAGCATAACTGAGTTTTGGAAGTTCATGGAA
GGACCCCTCATCGACGGCTTGTACTGGGACTCGTGGTATGGCACCAAACAGCTGTACAGTGTGA
AGAACAGCAGCCGCATCTACTACGAGAACGTTCTTCTCGGCATCCCCAGAGTGCGGCAACTGCG
AGTCCGAAACAACACTTGCAAGGTCTACCCAGCTTTCAGTCCCTGGTCAGCGACTGCTACAGC
AAGTACACAGTGGAAAACGAAGACTTCTCTGATTTTGGCCTCAAACGCAATCCAGAATGGACGC
ACACGCCTTCTTCCCGCACTGCCCCATGGCACTGGGGGTTTGTGGCGTATACCGAGATGGAGG
ATATATAGTCACTTATCAAAATCAAAATCTGAAACCAAAGCCAAATTTGTTGACCTTCGACTG
AACAACCTGGATTAGCAGAGGCACCAGGGCTGTTTTTATTGATTTCTCCCTGTACAATGCTAATG
TCAACCTGTTTTTGCATCATCAGGCTGCTGGCAGAGTTCCCTGCGACGGGTGGGCTCCTCACCTC
CTGGCAGTTCTACTCTGTGAAGCTCCTCAGATACGTCTCCTACTACGACTACTTCATTGCCTCC
TGTGAAGTCATATTTTGTATTTTCTCTTTGTCTTCATAATACAAGAACTGAGGAAAGTGAACG
AGTTTAACTCTGCCTATTTTCAAGAGTGTCTGGAACCTGGCTGGAGATGCTGCTCCTGCTGCTCTG
TTTTCTCGCCGTGTCTTTCTATGCATACTGTAACATGCAGAGCTTTCTCTTGCTTGGACAGCTG
CTGAAAAACACTGACAGCTATCCCGACTTTTACTTCCTTGCATACTGGCACATTTACTATAACA
ACGTAATTGCTATCACTATCTTCTTTGCATGGATAAAGATATTCAAGTTCATAAGCTTCAATGA
GACAATGTCGCAGCTGTCATCAACACTCTCCCGCTGCATGAAGGACATCGTGGGGTTTCGCCATC
ATGTTCTTCATCATCTTCTCTGCTTATGCCCAGTTGGGATTTCTGGTTTTTGGGTCACAGGTTG
ATGATTTTTCAACTTTTCAAAATTCATATTTGCACAATTTTGAATTGTCCTCGGGGACTTTAA
CTTTGCTGGCATCCAGCAGGCCAACTGGATCTTGGGGCCCATCTACTTCATCACGTTTCATCTTC
TTTGTGTTCTTTGTGCTCCTGAACATGTTCTTGGCAATAATTAATGACACCTATTCTGAAGTTA
AGGCTGATTATTCAATAGGCAGAAGACCAGATTTTGAACCTGGTAAAATAATTCAAAGAGTTG
CTTTAATGTTCTCGAGAACTCAGACTCAAGAAAGCTCAAGCTAAAGAAGAAAAGAAAATGCAA
ACCACTGACTTGGCCAGAGAGCCAGAAGAGAAGGCTTTGATGAAAGTGAGATCCAAGAGGCAG
AGCAGATGAAAAGATGGAAGGAAAGGCTTGAAAAAAGTATTATTCTACAGAAATTCAAGACGA
TTATCAGCCTGTCACTCAGCAAGAATTCCGAGAGCTCTTTTTATACGCGGTGGAGCTTGAGAAG
GAATTACACTATGTCAGTTTAAACTGAACCAACTGATGAGAAAGCTGCACTAGCAGGCTGACA
AGTGGAGTCATTTTATAAGAAAAGGCAACCGAAGAATTTTATTTCAGTATGAAGATTTTTTCCTC
TCAATTTTCTTCAGTAATGCAAAAGAAAAACCAAAAAGTAACCAGAAGTGCTTTTTATTTCAAA
GTTCTTGAAGTAAAGAGTAAACTCTTGTCCTTTGCTAACAGCCGTGTCTGCAGTAAACAAT
GAAGGAGCCTGCGTGTTCCTAAGTGTGGAGAGGATCTGCGGGAATGTGGAACAGCTTTCCTTG
CCTACTGGAACCAACAAACAAGCACACAATGGGACTCTCTGAGTGCCTGACAAAGTGAACGCAAG
TACAGCCAAGCACACATGGTGAACCTGTCAGGGAACACAAGCACTTTATGGCGTCACTTTCAAG
GAACATATTTTATATGGATTTTGAAGAGTCTTGTGTTGCTGATAAGAACTTCAAGAAGTCTAAGC
TTGGCTTTGATTCTCTTGATTCTTATATTCCTCAAGCACCGGAACACGATCCTCCTTCTGGG
CATTCTAGGGAAGATAAACTCTGTAAAGCAAAAAAAGAAAAAAGAAAAA (SEQ ID
NO: 1)

FIGURE 1

MSEATWWYRGGTSKHDHLHYRREAENVNTTLEELLLYFIFLINLCILTFGMVNPHEMYLKNKMSSL
FVDTSLPDDERSSSFRSIRSITEFWKFMEGPLIDGLYWDWYGTKQLYSVKNSRIYYENVLLGI
PRVRQLRVRNNTCKVYPAFQSLVSDCYSKYTVENEDFSDFGLKRNPEWTHTPSSRTAPWHWGFV
GVYRDGGYIVTLSSKSKSETKAKFVDLRLNNWISRGTRAVFIDFSLYNANVNLFCIIRLLAEFPA
TGGLTTSWQFYSVKLLRYVSYYDYFIASCEVIFCIFLFVFI IQELRKVNEFKSAYFRSVWNWLE
MLLLLLCFLAVSFYAYCNMQSFLLLGQLLKNTDSYPDFYFLAYWHIYYNNVIAITIFFAWIKIF
KFISFNETMSQLSSTLSRCMKDIVGFAIMFFIIFSAYAQLGFLVFGSQVDDFSTFQNSIFAQFR
IVLGDFNFAGIQQANWILGPIYFITFIFFVFFVLLNMFLAIINDTYSEVKADYSIGRRPDFELG
KIIQKSCFNVLEKLRLKKAQAKEEKKMQTTDLAQRARREGFDESEIQEAEQMKRWKERLEKKYY
STEIQDDYQPVTQQEFRELFYAVELEKELHYVSLKLNQLMRKLH (SEQ ID NO:2)

FIGURE 2

underlined = deleted in targeting construct

BOLD = sequence flanking Neo insert in targeting construct

GGCAGAACGGGCATAGTGGAAACGGACCGTGTGGTATGCTCCGCGATGTCTGAGGCGACTT
GGTGGTACCGAGGAGGGACTTCAAAACATGACCTGCATTACAGAAGGGAAGCGGAGGTTA
ACACCACACTCGAGGAGTTGTTACTCTATTTTATTTTCTTAATAAATCTATGCATATTGA
CTTTTGGGATGGTCAATCCACATATGTACTATTTAAACAAAGTTATGTCTCTCTGTTTG
TGGACACTTCTCTACCTGATGATGAAAGAAGCAGCTTTAGGTCCATTCCGAGCATAACTG
AGTTTTGGAAGTTCATGGAAGGACCCCTCATCGACGGCTTGTACTGGGACTCGTGGTATG
GCACCAAACAGCTGTACAGTGTGAAGAACAGCAGCCGCATCTACTACGAGAACGTTCTTC
TCGGCATCCCCAGAGTGCGGCAACTGCGAGTCCGAAACAACACTTGCAAGGTCTACCCAG
CTTTCCAGTCCCTGGTCAGCGACTGCTACGCAAGTACACAGTGGAAAACGAAGACTTCT
CTGATTTTGGCCTCAAACGCAATCCAGAATGGACGCACACGCCTTCTTCCGCACTGCC
CATGGCACTGGGGGTTTGTGGCGTATACCGAGATGGAGGATATATAGTCACGTTATCAA
AATCAAAATCTGAAACCAAAGCCAAATTTGTTGACCTTCGACTGAACAACCTGGATTAGCA
GAGGCACCAGGGCTGTTTTTATTGATTTCTCCCTGTACAATGCTAATGTCAACCTGTTTT
GCATCATCAGGCTGCTGGCAGAGTTCCTGCGACGGGTGGGCTCCTCACCTCCTGGCAGT
TCTACTCTGTGAAGCTCCTCAGATACGTCTCCTACTACGACTACTTCATTGCCTCCTGTG
AAGTCATATTTTGTATTTTCTCTTTGTCTTCATAATACAAGAACTGAGGAAAGTGAACG
AGTTTAAGTCTGCCTATTTTCAGAAAGTGTCTGGAAGTGGCTGGAGATGCTGCTCCTGCTGC
TCTGTTTTCTCGCCGTGTCTTTCTATGCATACTGTAACATGCAGAGCTTTCTCTTGCTTG
GACAGCTGCTGAAAAACACTGACAGCTATCCCGACTTTTACTTCCCTTGCCATACTGGCACA
TTTACTATAACAACGTAATTTCTATCACTATCTTCTTTGCATGGATAAAGATATTCAAGT
TCATAAGCTTCAATGAGACATGTGCGAGCTGTCAACACTCTCCCGCTGCATGAAGG
ACATCGTGGGGTTCGCCATCATGTTCTTCATCATCTTCTCTGCTTATGCCAGTTGGGAT
TTCTGGTTTTTGGGTCAAGGTTGATGATTTTCAACTTTTCAAAATTCATATTTGCAC
AATTTCAATTGTCTCTCGGGACTTTAACTTTGCTGGCATCCAGCAGGCCAACTGGATCT
TGGGGCCCATCTACTTATCACGTTTCTCTTTGTGTTCTTTGTGCTCCTGAACATGT
TCTTGGCAATAATTAATGACACCTATTCTGAAGTTAAGGCTGATTATTCAATAGGCAGAA
GACCAGATTTTGAAGTTGGTAAAATAATTCAAAAGAGTTGCTTTAATGTTCTCGAGAAAC
TCAGACTCAAGAAAGCTCAAGCTAAAGAAGAAAAGAAAATGCAAACCACTGACTTGGCCC
AGAGAGCCAGAAGAGAAGGCTTTGATGAAAGTGAGATCCAAGAGGCAGAGCAGATGAAAA
GATGGAAGGAAAGGCTTGAAAAAAGTATTATTCTACAGAAATTCAAGACGATTATCAGC
CTGTCACTCAGCAAGAATTCGAGAGCTCTTTTTATACGCGGTGGAGCTTGAGAAGGAAT
TACACTATGTGAGTTTAAACTGAACCAACTGATGAGAAAGCTGCACTAGCAGGCTGACA
AGTGGAGTCATTTTATAAGAAAAGGCAACCGAAGAATTTCAATCAGTATGAAGATTTTTT
CCTCTCAATTTTCTTCAGTAATGCAAAAGAAAAACCAAAAGTAACCAGAAGTGCTTTTT
ATTTCAAAATTTCTTGAAGTAAAGAGTAAACTCTTGTCTTTGCTAACAGCCGTGCTG
CAGTAAACCAATGAAGGAGCCTGCGTGTTCCTAAGTGTGGAGAGGATCTGCGGAATGT
GGAACACTTTCTTGCCTACTGGAACCACAAACAAGCACACAATGGGACTCTCTGAGTG
CCTGACAAAGTGAACGCAAGTACAGCCAAGCACACATGGTGAAGTGTGAGGAAACACAAG
CACTTTATGGCGTCAACTTTCAAGGAACATATTTTATATGGATTTTGAAGAGTCTTGTTT
GCTGATAAGAACTTCAAGAAGTCTAAGCTTGGCTTTGATTCTCTGTATTCCTTATATTC
CTCAAGCACCGGAACACGATCCTCCTTCTGGGCATTCTAGGGAAGATAAACTCTGTAA
AGCAAAAAAAAAAGAAAAAAAAAAAAAAAA

FIGURE 3

Gene Sequence Structure

*

211 bp

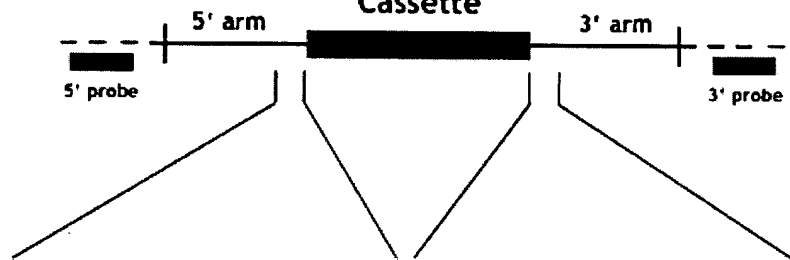
Sequence Deleted

355 bp

Size of full-length
cDNA: 2486 bp

Targeting Vector*
(genomic sequence)

LacZ-Neo
Cassette



Arm Length:
5': 2.5 kb
3': 1 kb

———— Targeting Vector
----- Endogenous Locus

* Not drawn to scale

```

5' >CATGCATACATGGGTGCACGT
GCGTGCGTGCAAACACCCACACAC
AGATAGATAAATGTAAAAAATAC
TTAGGCAATGCTGTTTTTTTAAAG
CTTGGGAAATTAATATTTTTGAAC
ACTGCTGATGTGGGATTTTAAGTT
ACAAATGTTAACTTGTTCTCCTT
AGTGACTTTTGGGATGGTCAATCC
ACATATGTACT<3'
(SEQ ID NO:3)
  
```

```

5' >GTATGGCAACAAACAGCTGTA
CAGTGTGAAGAACAGCAGCCGCAT
CTACTACGAGAACGTTCTTCTCGG
CATCCCAGAGTGCGGCAACTGCG
AGTCCGAAACAACACTTGCAAGGT
CTACCCAGCTTTCAGTCCCTGGT
CAGCGACTGCTACAGCAAGTACAC
AGTGGAACGAAAGACTTCTCTGA
TTTTGGCCTCA<3'
(SEQ ID NO:4)
  
```

FIGURE 4